E-crime

CS642: Computer Security

Professor Ristenpart

http://www.cs.wisc.edu/~rist/

rist at cs dot wisc dot edu
ATTN; Thomas Ristenpart

★ Alan Daniel <alandaniel33@yahoo.co.uk> to crypto2015@iacr.org

Confidant: Gregor Leander

I am Barr. Allan Daniel (retired).

I have a knowledge of an unclaimed estate proceed left in a codicil by a deceased Veteran, Scott G. Ristenpart....{2nd Battalion, 505th Parachute Infantry Regiment, 3rd Brigade Combat Team, 82nd Airborne Division at Ft Bragg NC...}.

I am contacting you to seek for your co-operation to stand in as a possible heir since you have the same last name as the deceased beneficiary so that the proceed of this estate could be taken as a fortune before the Probate ref: no is revoked for a lack of "Executor/Executrix".

If you agree to this confidential consultation....I will advise on the programmed of work to expedite this claim through a legitimate arrangement for a mutual benefit..

Thank you for your anticipated response.

Best regards,

Allan Daniel Jr.
Spam, phishing, scams

• Spam
  – unsolicited bulk emails
  – 2006: 80% of emails on web, 85 billion messages a day
  – 2009: 95% of emails blocked as spam (100s of billions)

• Scam spam
  – Nigerian emails (advanced fee fraud / confidence trick)

• Phishing
  – trick users into downloading malware, submitting CC info to attacker, etc.
  – Spear phishing: targeted on individuals (used in high-profile intrusions)
Spanish Prisoner confidence trick

- 19th century
- In contact with rich guy in Spanish prison
- Just need a little money to bribe guards, he’ll reward you greatly
Hi Dear,

I am Mrs. Zarina Al-Usman, I have been diagnosed with Esophageal cancer. It has defied all forms of medical treatment, and Right now, I have only about a few months to live and I want you to Distribute my funds worth Twelve Million Five Hundred Thousand US Dollars to charities homes in your country.

I have set aside 40% for you and your family so keep this as a secret to yourself because this will be my last wish.
Yours Truly,

Mrs. Zarina Al-Usman

WebMail  FDV - MG
Faculdade Viçosa
This is an automatic notification of your current disk space usage on the CSE mail server:

```
csemailbox.ucsd.edu
```

**Your account status:**

- Current utilization: 95.33%
- Space used: 976 MB
- Available space: 47 MB
- Account limit: 1024 MB

Once your quota has been reached, mail will no longer be delivered to your account, and will be returned to the sender as undeliverable.

If you are not sure where to look for mail that can likely be deleted to clear space in your account, you may likely have large amounts of mail in your Trash and/or Junk folders. Also, you may have a large amount of mail accumulating in your Sent folder over time, if you have configured your mail client to automatically save sent messages.

Your account limit may be increased for an additional charge, as per the CSE Recharge Policy. Please contact CSEHelp regarding quota increases.

Please reply to this message or contact CSEHelp <csehelp@cs.ucsd.edu> if you have any questions or require assistance.

Thank you,
Spam volume

Billions of messages

http://www.senderbase.org/static/spam/#tab=1
Spam

• The frontend (email recipients)
  – Filtering, classification
  – Psychology

• The backend (email generation)
  – Open email relays
  – Botnets
  – Social structure
    • Affiliates
    • Criminal organizations
Spam Classifiers

This classifier will be trained from a large corpus of labeled data.

Spam / Ham

Ham

Spam
Naïve Bayes Classifier

Represent email as “bag of words”

\[
\text{quota} \quad 1 \quad x_1 \\
\text{webmail} \quad 4 \quad x_2 \\
\text{wisconsin} \quad 0 \quad x_3 \\
\text{viagra} \quad 0 \quad x_4 \\
\vdots \quad \vdots
\]

Intuition: spam and ham have different distribution of keywords

\[
\Pr[\text{spam} | x_1, x_2, \ldots, x_n] = \frac{\Pr[x_1, x_2, \ldots, x_n | \text{spam}] \Pr[\text{spam}]}{\Pr[x_1, x_2, \ldots, x_n]}
\]

Bayes’ theorem

\[
= \frac{\Pr[\text{spam}] \prod \Pr[x_i | \text{spam}]}{\Pr[x_1, x_2, \ldots, x_n]}
\]

“Naïve”: assume words independent

\[
\Pr[\text{ham} | x_1, x_2, \ldots, x_n] = \frac{\Pr[\text{ham}] \prod \Pr[x_i | \text{ham}]}{\Pr[x_1, x_2, \ldots, x_n]}
\]
Naïve Bayes Classifier

Represent email as “bag of words”

- quota 1 \( x_1 \)
- webmail 4 \( x_2 \)
- wisconsin 0 \( x_3 \)
- viagra 0 \( x_4 \)
- \( \vdots \) \( \vdots \)

Intuition: spam and ham have different distribution of keywords

\[
\Pr(\text{spam} | x_1, x_2, \ldots, x_n) = \frac{\Pr(\text{spam}) \prod \Pr(x_i | \text{spam})}{\Pr(x_1, x_2, \ldots, x_n)}
\]

\[
\Pr(\text{ham} | x_1, x_2, \ldots, x_n) = \frac{\Pr(\text{ham}) \prod \Pr(x_i | \text{ham})}{\Pr(x_1, x_2, \ldots, x_n)}
\]

Classify as spam if: \( \Pr(\text{spam} | x_1, x_2, \ldots, x_n) > \Pr(\text{ham} | x_1, x_2, \ldots, x_n) \)
Naïve Bayes Classifier

Represent email as “bag of words”

\[
\text{Pr[ spam | } x_1, x_2, \ldots, x_n ] = \frac{\text{Pr[spam] } \prod \text{Pr}[x_i | \text{spam} ]}{\text{Pr}[x_1, x_2, \ldots, x_n]}
\]

\[
\text{Pr[ ham | } x_1, x_2, \ldots, x_n ] = \frac{\text{Pr[ham] } \prod \text{Pr}[x_i | \text{ham} ]}{\text{Pr}[x_1, x_2, \ldots, x_n]}
\]

Classify as spam if:

\[
\text{Pr[spam] } \prod \text{Pr}[x_i | \text{spam} ] > \text{Pr[ham] } \prod \text{Pr}[x_i | \text{ham}]
\]

Intuition: spam and ham have different distribution of keywords

Estimate these from labeled training data
Spam classifiers

- Real classifiers more complex than this
  - Other features: Who is sender? How many links embedded? Is it from an open mail relay?
  - Can update in real-time given labelings by user
  - For larger orgs, can leverage wide view across many email recipients

- Nowadays some companies do pretty good job of making sure spam doesn’t hit your inbox
  - 95% of email gets filtered as spam (2009, ENISA Spam Survey)
Spam

• The frontend (email recipients)
  – Filtering, classification
  – Psychology, usability

• The backend (email generation)
  – Open email relays
  – Botnets
  – Social structure
    • Affiliates
    • Criminal organizations
Botnets

- Botnets:
  - Command and Control (C&C)
  - Zombie hosts (bots)
- C&C type:
  - centralized, peer-to-peer
- Infection vector:
  - spam, random/targeted scanning
- Usage:
  - What they do: spam, DDoS, SEO, traffic generation, ...

Figure 1: The Storm botnet hierarchy.
How to make money off a botnet?

- **Rental**
  - “Pay me money, and I’ll let you use my botnet... no questions asked”

- **DDoS extortion**
  - “Pay me or I take your legitimate business off web”

- **Bulk traffic selling**
  - “Pay me to direct bots to websites to boost visit counts”

- **Click fraud, SEO**
  - “Simulate clicks on advertised links to generate revenue”
    - Cloaking, link farms, etc.

- **Theft of monetizable data** (eg., financial accounts)

- **Data ransom**
  - “I’ve encrypted your harddrive, now pay me money to unencrypt it”

- **Advertise products**
## Underground forums

<table>
<thead>
<tr>
<th>Category</th>
<th>Threads</th>
<th></th>
<th>Users</th>
<th></th>
<th>Top Subcategory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>S</td>
<td>B</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>payments</td>
<td>8,507</td>
<td>8,092</td>
<td>1,539</td>
<td>1,409</td>
<td>paysafecard</td>
</tr>
<tr>
<td>game-related</td>
<td>2,379</td>
<td>2,584</td>
<td>924</td>
<td>987</td>
<td>steam</td>
</tr>
<tr>
<td>accounts</td>
<td>2,119</td>
<td>2,067</td>
<td>850</td>
<td>974</td>
<td>rapidshare</td>
</tr>
<tr>
<td>credit cards</td>
<td>996</td>
<td>1160</td>
<td>467</td>
<td>566</td>
<td>unspecified cc</td>
</tr>
<tr>
<td>software/keys</td>
<td>729</td>
<td>1410</td>
<td>422</td>
<td>740</td>
<td>key/serial</td>
</tr>
<tr>
<td>fraud tools</td>
<td>652</td>
<td>1155</td>
<td>363</td>
<td>601</td>
<td>socks</td>
</tr>
<tr>
<td>tutorials/guides</td>
<td>950</td>
<td>537</td>
<td>562</td>
<td>393</td>
<td>tutorials</td>
</tr>
<tr>
<td>mail/drop srvs</td>
<td>751</td>
<td>681</td>
<td>407</td>
<td>364</td>
<td>packstation</td>
</tr>
<tr>
<td>merchandise</td>
<td>493</td>
<td>721</td>
<td>264</td>
<td>404</td>
<td>ipod</td>
</tr>
<tr>
<td>services</td>
<td>266</td>
<td>916</td>
<td>176</td>
<td>555</td>
<td>carder</td>
</tr>
</tbody>
</table>

Table 6: Top 10 most commonly traded merchandise categories on LC.

How to make money off financial credentials?

• Money mules
  – Deposits into mules’ account from the victim’s
  – Mule purchases items using stolen CCN, sells them online
  – Mule withdraws cash from ATMs using victim credentials

• Wires money to (frequently) former Soviet Union
Dear Student,

I would like to offer you a new interesting and respectable job! We are looking for people to work as professional distance-based typists. No experience is needed! If you’re eager to use your skills to make some additional cash, then you might want to consider a home typing position!

All data entry operators work from home and are independent contractors. You typically set your own hours and work from home on projects that are enjoyable! Average monthly earnings start from $1000 to $3000 or more.

Requirements:
- Computer with Internet access.
- Good Typing Skills.
- Basic Internet knowledge.
- Basic Computer and Typing Skills.

You will not have to devote full-time hours. These assignments can be done on your time. They may be done in Internet cafes or wherever you can get Internet access!

If you are interested just reply to my email!

Best Regards,

Richard Hill
Local Recruitment Manager
Organized cyber criminals stole almost $11 million in two highly coordinated ATM heists in the final days of 2012, KrebsOnSecurity has learned. The events prompted Visa to warn U.S. payment card issuers to be on high-alert for additional ATM cash-out fraud schemes in the New Year.
How to make money off a botnet?

• Rental
  – “Pay me money, and I’ll let you use my botnet... no questions asked”

• DDoS extortion
  – “Pay me or I take your legitimate business off web”

• Bulk traffic selling
  – “Pay me to direct bots to websites to boost visit counts”

• Click fraud, SEO
  – “Simulate clicks on advertised links to generate revenue”
  – Cloaking, link farms, etc.

• Theft of monetizable data (eg., financial accounts)

• Data ransom
  – “I’ve encrypted your harddrive, now pay me money to unencrypt it”

• Advertise products
Botnets

• Botnets:
  – Command and Control (C&C)
  – Zombie hosts (bots)

• C&C type:
  – centralized, peer-to-peer

• Infection vector:
  – spam, random/targeted scanning, drive-by exploit

• Usage:
  – What they do: spam, DDoS, SEO, traffic generation, ...

Figure 1: The Storm botnet hierarchy.
Agobot (circa 2002)

- IRC botnet
- Rich feature set:
  - Well-documented, modular codebase
  - IRC-based C&C system
  - Large catalogue of remote exploits
  - Limited code obfuscation and anti-disassembly techniques
  - Built-in data collection
  - Mechanisms to disable antivirus
  - Large set of bot commands
Storm botnet (2007-08)

- Sept 2007
  - Media: 1 – 50 million bots
  - More likely: 10,000s to 100,000s
- Early spam campaigns used titles such as “230 dead as storm batters Europe.”
- Propagated via spam linking to malware
- Thought to be controlled by Russian Business Network

Features:
- Uses P2P (Overnet/Kademlia)
- Uses fast-flux DNS for hosting on named sites
- Binary has gone through many revisions
- Features of P2P network have evolved with time
- Hides on machine with rootkit technology

Enright 2007

See also: http://cseweb.ucsd.edu/~savage/papers/login08.pdf
Overnet DHT

DHT stands for Distributed Hash Table

Peer-to-peer protocol for storing and retrieving data.

Storm uses DHT to store IP addresses of proxies

Proxies maintain contact with Bot C&C servers

Publicly addressable bots advertise as able to C&C proxy

Figure 1: The Storm botnet hierarchy.
Fast-flux DNS

Spam campaign that directs users to pharmashop.com

Single flux:
• Change A record for pharmashop.com quickly to point to different compromised systems
• Short TTL (e.g., 5 minutes)

Double flux:
• Change NS record for pharmashop.com to point to different compromised systems

Similar to round-robin DNS as used by major websites
Holz et al. 2008
http://pi1.informatik.uni-mannheim.de/filepool/publications/fast-flux-ndss08.pdf
Studying fast-flux

CDF of average time between DNS A record address changes

Konte et al. 2008
http://www.cc.gatech.edu/~feamster/papers/fastflux-tr08.pdf
Studying fast-flux

<table>
<thead>
<tr>
<th>type</th>
<th>Minutes/IP</th>
<th>IP/Day</th>
<th>A-TTL</th>
<th>NS-TTL</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td>73.55</td>
<td>55.90</td>
<td>1832.84</td>
<td>37348.75</td>
</tr>
<tr>
<td>max</td>
<td>634.50</td>
<td>261.54</td>
<td>21598.03</td>
<td>65535.00</td>
</tr>
<tr>
<td>min</td>
<td>5.51</td>
<td>2.27</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Xu et al. 2013
These techniques may already account for wide discrepancies in the estimated size of various botnets seen in the media. With so many groups taking uncoordinated actions with noticeable effects, it is only a matter of time before problems occur. For example, one possible problem would be the effect of a researcher inflating the perceived size of a botnet that is the subject of a criminal investigation. If such a case resulted in a successful prosecution, and a damage estimate were to be derived based on the inflated count of "infected" hosts multiplied by some estimated cost of cleanup accepted by the courts, the resulting damages would be similarly inflated. This is not out of the question, as several cases in the past few years have included evidence obtained by law enforcement agents as to the number of bots under the control of the suspect. It is likely that some of these suspects, even if they admit to the numbers stated, may not know precisely how many hosts they truly did compromise and control.

One final interesting observation, which we have not seen noted in any other research to date, are the downward spikes in the bottom line (the reachable and responsive peers) of the Naguche botnet.

Geolocating bots enumerated for Naguche botnet
Dittrich and Dietrich, “Discovery Techniques for P2P Botnets”
How can we measure botnets?

- Monitor C&C traffic
  - Join IRC channels
  - Network captures at ISPs
- DNS interrogation
  - Identify domain names that point to bots (FastFlux)
- P2P bot crawlers
  - Crawl overlay network to enumerate hosts
- Botnet infiltration
  - Take over C&C
  - Participate as proxies, bots, etc.
## Botnet measurement methods

<table>
<thead>
<tr>
<th>Technique</th>
<th>Description</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitor endpoint</td>
<td>monitor traffic of a bot</td>
<td>simple, generally applicable</td>
<td>limited view, encryption</td>
</tr>
<tr>
<td>Internet telescopes</td>
<td>monitor random-scan infection attempts</td>
<td>botnet-wide view</td>
<td>limited applicability</td>
</tr>
<tr>
<td>Monitor IRC</td>
<td>record IRC C&amp;C traffic</td>
<td>simple, botnet-wide view</td>
<td>only IRC botnets</td>
</tr>
<tr>
<td>DNS redirect</td>
<td>hijack C&amp;C via DNS</td>
<td>measure infection size</td>
<td>limited applicability</td>
</tr>
<tr>
<td>Sybil monitoring</td>
<td>monitor numerous bots</td>
<td>simple, passive</td>
<td>resource-intensive, limited view, structured P2P</td>
</tr>
<tr>
<td>Botnet crawling</td>
<td>crawl botnet overlay</td>
<td>enumerate large portion of botnet</td>
<td>detectable</td>
</tr>
<tr>
<td>DNS cache probing</td>
<td>probe DNS caches for botnet C&amp;C</td>
<td>simple, passive</td>
<td>loose lower-bound</td>
</tr>
<tr>
<td>DNSBL counter-intelligence</td>
<td>sniff DNSBL traffic, heuristically identify bots</td>
<td>passive</td>
<td>limited applicability</td>
</tr>
<tr>
<td>Flow analysis</td>
<td>detect botnets via flow-based anomaly detection</td>
<td>wide-scale, handles encryption</td>
<td>tailored to IRC botnets</td>
</tr>
</tbody>
</table>
### Size estimates from literature as of 2008

<table>
<thead>
<tr>
<th>Study</th>
<th>Method(s) used</th>
<th>C&amp;C’s observed</th>
<th>Largest botnet size</th>
<th>Total # of infected hosts</th>
</tr>
</thead>
<tbody>
<tr>
<td>[13]</td>
<td>IRC monitoring</td>
<td>~100</td>
<td>226,585</td>
<td>–</td>
</tr>
<tr>
<td>[8]</td>
<td>IRC monitoring</td>
<td>~180</td>
<td>~50,000</td>
<td>~300,000</td>
</tr>
<tr>
<td>[22]</td>
<td>DNS cache probing</td>
<td>65</td>
<td>–</td>
<td>85,000</td>
</tr>
<tr>
<td></td>
<td>IRC monitoring</td>
<td>&gt;100</td>
<td>&gt;15,000</td>
<td>~3,000</td>
</tr>
<tr>
<td>[23]</td>
<td>DNS cache probing</td>
<td>100</td>
<td>–</td>
<td>88,000</td>
</tr>
<tr>
<td></td>
<td>IRC monitoring</td>
<td>472</td>
<td>~100,000</td>
<td>426,279</td>
</tr>
<tr>
<td>[5]</td>
<td>DNS redirection</td>
<td>~50</td>
<td>&gt;350,000</td>
<td>–</td>
</tr>
<tr>
<td>[15]</td>
<td>flow analysis</td>
<td>~376</td>
<td>–</td>
<td>~6,000,000</td>
</tr>
<tr>
<td>[7]</td>
<td>botnet crawling</td>
<td>1</td>
<td>~160,000</td>
<td>~44,000</td>
</tr>
</tbody>
</table>

Figure 2: Size estimates from the literature. All sizes are the maximum ones given in the appropriate study and the final column represents the total number of infected hosts over all botnets encountered.
Size of Storm botnet

The blue peers count is all peers being probed at a time. This includes live, active, dead, and unknown states. The peers line is not the size of the network. The active line is much closer to the instantaneous size of the network.

It can be seen in the month and year chart that Microsoft made a measurable dent in the network with the MRT Storm (Nuwar) release.
Botnet infiltration studies

• Spamalytics (Kanich et al., 2008)
  – Storm botnet
  – Rewrote spam to redirect to researcher-controlled websites
  – **Goal**: click-through rate measurement
Kanich et al., Spamalytics: An Empirical Analysis of Spam Marketing Conversion, 2008
The victims

Figure 9: Geographic locations of the hosts that “convert” on spam: the 541 hosts that execute the emulated self-propagation program (light grey), and the 28 hosts that visit the purchase page of the emulated pharmacy site (black).

Kanich et al., Spamalytics: An Empirical Analysis of Spam Marketing Conversion, 2008
Observed Conversion Rate

• 350 million email messages delivered
• 26 day campaign
• 28 “sales”
  – 0.00001%
  – 27 of these male-enhancement products
Botnet takeover studies

• Spamalytics (Kanich et al., 2008)
  – Storm botnet
  – Rewrote spam to redirect to researcher-controlled websites
  – **Goal**: click-through rate measurement

• Torpig C&C sinkholing (Stone-gross et al., 2009)
  – Torpig botnet
  – Setup researcher controlled C&C server (DNS fastflux)
  – **Goal**: analysis of stolen data
In fact, the injected content carefully reproduces card numbers and social security numbers, in a form that asks the user for sensitive information, for example, credit the user's browser to. This content typically consists of an HTML URL from the injection server and injects the returned content into visits the trigger page. Torpig requests the injection of times it can be launched. The second step occurs when the user attack tests whether the attack is active and the maximum number URL page where the attack should be triggered we call this page the server. Cally a banking web site, Torpig issues a request to an attacks occur in two steps. First, whenever the infected machine served during the passive monitoring it normally performs. These data we call this reply an injection mechanism defies all phishing indicators included in the style and look-and-feel of the target web site. Furthermore, the SSL configuration appears modern browsers. For example, the SSL configuration appears to perform "many-in-they-middle" attacksx, and send the stolen information to its 'r' server. This communication is also over HTTP and is protected by the data stolen since the previous reporting time. This communication in the Torpig botnet, except those used by the Mebroot structure to infect machines, retrieve updates, perform active phishing attacks, and send the stolen information to its 'r' server. It specifies how often the bot should contact the 'r' server, a set of parameters to fine tune the configuration file is obfuscated using a simple XOR encoding. This scheme was broken. In addition, the 'r' server can send a configuration file to the bot we call this reply an attack to. The command line interpreter looks, Eudora, instant messengers, Skype, ICQ, and system clients. FTP, LeechFTP, email clients, Thunderbird, Outlook, the file manager, and other popular applications such as web modules DLLs into a number of applications. These applications comprise the Torpig malware. Mebroot injects these

![Figure 1: The Torpig network infrastructure. Shaded in gray are the components for which a domain generation algorithm is used.](image)

**Figure 2: A man-in-the-browser phishing attack.**

Stone-Gross et al., Your Botnet is My Botnet: Analysis of a Botnet Takeover, 2009
Torpig relies on a fairly complex network infrastructure to infect machines, retrieve updates, perform communications with the server also occurs over HTTP, and the injection mechanism decreases the target webpage's style and "look and feel," without having to contact the C&C server. It also operates via HTTP requests and responses using a sophisticated encryption algorithm.

The C&C server can send a configuration (that is, the type and version number of the currently installed modules) and to receive any updates. All communication with the C&C server occurs periodically, in two-hour intervals, to report its current activity. Step 7: when browsing a targeted site, victim is redirected to a set of hard-coded servers to be used as backup, and the injection mechanism deorator (typically a banking webpage), Torpig issues a request to launch it. The second step occurs when the user visits the trigger page: Torpig requests the injection URL from the injection server and puts the returned content into the user's browser (step 7). This content is how often the bot should contact the C&C server, and the injection mechanism decorlates the modules and names them after existing, custom encryption algorithm.

Every 20 minutes, Torpig contacts the Torpig C&C server to upload stolen data (step 6). This communication with the C&C server is the component that we "hijacked." Step 1: attackers modify the injection mechanism decorator to infect machines, retrieve updates, perform communications with the server also occurs over HTTP, and the injection mechanism decreases the target webpage's style and "look and feel," without having to contact the C&C server. It also operates via HTTP requests and responses using a sophisticated encryption algorithm.

The C&C server can also send a configuration (that is, the type and version number of the currently installed modules) and to receive any updates. All communication with the C&C server occurs periodically, in two-hour intervals, to report its current activity. Step 7: when browsing a targeted site, victim is redirected to a set of hard-coded servers to be used as backup, and the injection mechanism decorlates the modules and names them after existing, custom encryption algorithm.

Even attentive users and the injection mechanism decorlates the modules and names them after existing, custom encryption algorithm.

The C&C server can also send a configuration (that is, the type and version number of the currently installed modules) and to receive any updates. All communication with the C&C server occurs periodically, in two-hour intervals, to report its current activity. Step 7: when browsing a targeted site, victim is redirected to a set of hard-coded servers to be used as backup, and the injection mechanism decorlates the modules and names them after existing, custom encryption algorithm.

Even attentive users and the injection mechanism decorlates the modules and names them after existing, custom encryption algorithm.
Figure 3. Unique bot IDs and IP addresses per hour. The number of unique IP addresses per hour provides a good estimation of Torpig’s live population.

Stone-Gross et al., Your Botnet is My Botnet: Analysis of a Botnet Takeover, 2009
Table 1. Data items sent to our C&C server by Torpig bots.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Data items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Form data</td>
<td>11,966,532</td>
</tr>
<tr>
<td>Email</td>
<td>1,258,862</td>
</tr>
<tr>
<td>Windows password</td>
<td>1,235,122</td>
</tr>
<tr>
<td>POP account</td>
<td>415,206</td>
</tr>
<tr>
<td>HTTP account</td>
<td>411,039</td>
</tr>
<tr>
<td>SMTP account</td>
<td>100,472</td>
</tr>
<tr>
<td>Mailbox account</td>
<td>54,090</td>
</tr>
<tr>
<td>FTP account</td>
<td>12,307</td>
</tr>
</tbody>
</table>

Stone-Gross et al., Your Botnet is My Botnet: Analysis of a Botnet Takeover, 2009
Table 3: Accounts at financial institutions stolen by Torpig.
Botnets

- Botnets:
  - Command and Control (C&C)
  - Zombie hosts (bots)
- C&C type:
  - centralized, peer-to-peer
- Infection vector:
  - spam, random/targeted scanning
- Usage:
  - What they do: spam, DDoS, SEO, traffic generation, ...
Botnet countermeasures?

- Infection prevention
- Infection detection
- C&C take-down
- Undermine the economics
  - Banking take-down
Infected detections & remediation

Anti-Botnet Efforts Still Nascent, But Groups Hopeful

Seven months after a government-industry coalition announced recommendations for ISPs to fight botnets, success is still a long way off

Nov 30, 2012 | 10:06 PM | 0 Comments
By Robert Lemos, Contributing Writer
Dark Reading
C&C takedowns

Microsoft Seizes ZeuS Servers in Anti-Botnet Rampage
BY KIM ZETTER 03.26.12  2:45 PM

It’s not the first time Microsoft has attempted to take down botnets. The company previously attacked three other botnets — Waledac, Rustock and Kelihos — through similar civil suits that allowed the company to seize web addresses and associated computers. The gains from such takedowns, however, are generally short-lived. After Waledac was targeted, the criminals behind it simply altered their software to thwart easy detection and launched a new botnet.

http://www.wired.com/threatlevel/2012/03/microsoft-botnet-takedown/
Botnet countermeasures?

• Infection prevention
• Infection detection
• C&C take-down
• Undermine the economics
  – Banking take-down
Studying grey/black market products

• Active measurement studies to:
  – Understand (probably illicit) services on web
  – Find ways to defuse underground markets

• Previous studies looked at botnets themselves and victims

• Let’s look at the “backend”
Spam-advertised products

- Pharmaceuticals
- Software
- Watches
- etc.

- What is order volume?
- What kinds of things are being purchased?
- What are weak links for disruption?

http://www.rioricopharmacy.com/
Measurement apparatus #1

Figure 6: How the purchase pair technique works

In this hypothetical situation, two measurement purchases are made that bracket some number of intervening purchases made by real customers. Because order number allocation is implemented by a serialized sequential increment, the difference in the order numbers between measurement purchases, $N = 23$, corresponds to the total number of orders processed by the affiliate program in the intervening time.

Proximate IP addresses and provided a unique email address for each order. We used five contact phone numbers for order confirmation, three from Google Voice and two via prepaid cell phones, with all inbound calls routed to the prepaid cell phones. In a few instances, it was necessary to place orders from IP addresses closely geographically located to the vicinity of the billing address for a given card, as the fraud check process for one affiliate program (EuroSofty) was sensitive to this feature. Another program (Royal Software) would only accept one order per IP address, requiring IP address diversity as well.

In total, we placed 60 such orders. We scheduled them both periodically over a three-week period as well as in patterns designed to help elucidate more detail about transaction volume and to test for internal consistency, as discussed below.

Finally, in addition to the raw data from our own purchase records, we were able to capture several purchase order numbers via forum scraping. This opportunity arose because affiliate programs typically sponsor online forums that establish a community among their affiliates and provide a channel for distributing operational information (e.g., changes in software or name servers, sharing experiences, which registrars will tolerate domains used to host pharmaceutical stores, and to raise complaints or questions). One forum in particular, for the GlavMed program, included an extended complaint thread in which individual affiliates complained about orders that had not yet cleared payment processing (important to them since affiliates are only paid for each settled transaction that they deliver). These affiliates chose to document their complaints by listing the order numbers they were waiting for, which we determined was in precisely the same format and numeric range as the order numbers presented to purchasers. By mining this forum, we obtained 677 numbers for past orders, including orders dating back to 2003.

<table>
<thead>
<tr>
<th>Affiliate Program</th>
<th>Phase 6</th>
<th>Phase 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>RX-Promotion</td>
<td>7465–6465</td>
<td>7466–7466</td>
</tr>
<tr>
<td>Pharmacy Express</td>
<td>8468–6868</td>
<td>8469–6969</td>
</tr>
<tr>
<td>GlavMed</td>
<td>6767–6969</td>
<td>6767–6969</td>
</tr>
<tr>
<td>Online Pharmacy</td>
<td>0460–6060</td>
<td>0461–6161</td>
</tr>
<tr>
<td>EvaPharmacy</td>
<td>6868–6969</td>
<td>6868–6969</td>
</tr>
<tr>
<td>88drugs</td>
<td>8868–6969</td>
<td>8868–6969</td>
</tr>
<tr>
<td>9RX</td>
<td>6565–6868</td>
<td>6566–6868</td>
</tr>
<tr>
<td>EuroSoft</td>
<td>8466–6666</td>
<td>8467–6767</td>
</tr>
<tr>
<td>Royal Software</td>
<td>7467–6767</td>
<td>7468–6868</td>
</tr>
<tr>
<td>SoftSales</td>
<td>7466–6666</td>
<td>7467–6767</td>
</tr>
</tbody>
</table>

Table 6: Active orders placed to sites of each affiliate program in the two different time phases of our study. In addition, we opportunistically gathered 677 orders for GlavMed covering the period between 2003 and 2006.

Note that this data contains an innate time bias since the date of complaint inevitably came a while later than the time of purchase (unlike our own purchases). For this reason, we identify opportunistically gathered points distinctly when analyzing the data. We will see below that the bias proves to be relatively minor.

We summarize the total data set in Table 6. It includes order numbers from 757 active purchases and 677 opportunistically gathered data points.

3.3 Consistency

While our initial observations of monotonicity are quite suggestive, we need to consider other possible explanations and confounding factors as well. Here we evaluate the data for internal consistency—the degree to which the data appears best explained by the sequential update hypothesis rather than other plausible explanations. At the end of the paper, we also consider the issue of external consistency using “ground truth” revenue data for one program.

Kanich et al., Show Me the Money: Characterizing Spam-advertised Revenue, 2011
Kanich et al., Show Me the Money: Characterizing Spam-advertised Revenue, 2011
Measurement Apparatus #2

Figure 6. How a user interacts with an EvaPharmacy Web site beginning with the landing page and then proceeding to a product page and the shopping cart. The main Web site contains embedded images hosted on separate compromised systems. When a browser visits such pages, the referrer information is sent to the image hosting servers for every new image visited. The product pages allow us to infer the selected product.

To quantify overall shopping cart addition activity, we compare the total number of visits to the number of visits to the shopping cart page. To quantify individual item popularity, we examine the subset of visits for which the customer workflow allows us to infer which specific item was added to the cart.

There are three key limitations to this approach. First and foremost, the final page in the purchasing workflow—the checkout page—generally does not include unique image content and thus does not appear in our logs. Even if it did, our approach could not determine whether checkout completed correctly. Thus, we can only observe that a user inserted an item into their cart but not that they completed a purchase attempt. In general, this is only an issue to the degree that shopping cart abandonment correlates with variables of interest. The second limitation is that pages typically use the same image for all dosages and quantities on a given product page, and therefore we cannot distinguish these features. We cannot distinguish between a user selecting 231 tablets of 36mg Viagra tablets vs. an order of 21 tablets each of 21mg. Finally, we cannot disambiguate multiple items selected for purchase. When a user visits a product page followed by the shopping cart page, we can infer that they selected the associated product. However, if the visitor then continues shopping and visits additional product pages, we cannot determine whether they added these products or simply examined them. Subsequent visits to the shopping cart page add few new recommended products: recommendations appear based on the first item in the cart. We choose the conservative approach and only consider the products that we are confident the user selected, which will cause us to underrepresent those drugs typically purchased together.

Another issue is that pharmacy formularies, while largely similar, are not identical between programs. Particularly, some pharmacy programs offer Schedule II drugs. However, since EvaPharmacy does not sell such drugs, our data does not capture this category of demand. Finally, our dataset also has potential bias due to the particular means used to drive traffic to it. We found that 56 of the 61 top landing pages observed in the hosting data also appeared in our spam-driven crawler database, demonstrating directly that these landing pages were advertised through email spam. While these pages could also be advertised using less risky methods such as SEO, this seems unlikely since spam-advertised URLs are swiftly blacklisted [25]. Thus, we suspect but cannot prove that our data may only capture the purchasing behavior for the spam-advertised pharmacies: different advertising vectors could conceivably attract different demographics with different purchasing patterns.

Given these limitations, we now report the results of two analyses: product popularity (what customers buy) and customer distribution (where the money comes from).

4.3 Product popularity

Our first analysis focuses on simple popularity, what individual items users put into their shopping carts and what broad, seller-defined categories of pharmaceuticals were popular during our measurement period. Although naturally dominated by the various ED and sexually-related pharmaceuticals, we find a surprisingly long tail: indeed, 49% of all items added to the cart were not in this category. We observed 390 distinct products, including popular mass-market products such as Zithromax (42), Compia (38), Nexium (37), and Propecia (38), but also Imo (22), a commonly prescribed antibiotic, Actos (7), a treatment for Type 3 diabetes, Uspar (23), an antianxiety, Seoquel (0), an anti-schizophrenia, Lomid (9), an ovulation inducer, and Gleevec (2), used to treat Leukemia and other cancers.
<table>
<thead>
<tr>
<th>Product</th>
<th>Quantity</th>
<th>Min order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generic Viagra</td>
<td>568</td>
<td>$78.80</td>
</tr>
<tr>
<td>Cialis</td>
<td>286</td>
<td>$78.00</td>
</tr>
<tr>
<td>Cialis/Viagra Combo Pack</td>
<td>172</td>
<td>$74.95</td>
</tr>
<tr>
<td>Viagra Super Active+</td>
<td>121</td>
<td>$134.80</td>
</tr>
<tr>
<td>Female (pink) Viagra</td>
<td>119</td>
<td>$44.00</td>
</tr>
<tr>
<td>Human Growth Hormone</td>
<td>104</td>
<td>$83.95</td>
</tr>
<tr>
<td>Soma (Carisoprodol)</td>
<td>99</td>
<td>$94.80</td>
</tr>
<tr>
<td>Viagra Professional</td>
<td>87</td>
<td>$139.80</td>
</tr>
<tr>
<td>Levitra</td>
<td>83</td>
<td>$100.80</td>
</tr>
<tr>
<td>Viagra Super Force</td>
<td>81</td>
<td>$88.80</td>
</tr>
<tr>
<td>Cialis Super Active+</td>
<td>72</td>
<td>$172.80</td>
</tr>
<tr>
<td>Amoxicillin</td>
<td>47</td>
<td>$35.40</td>
</tr>
<tr>
<td>Lipitor</td>
<td>38</td>
<td>$14.40</td>
</tr>
<tr>
<td>Ultram</td>
<td>38</td>
<td>$45.60</td>
</tr>
<tr>
<td>Tramadol</td>
<td>36</td>
<td>$82.80</td>
</tr>
<tr>
<td>Prozac</td>
<td>35</td>
<td>$19.50</td>
</tr>
<tr>
<td>Cialis Professional</td>
<td>33</td>
<td>$176.00</td>
</tr>
<tr>
<td>Retin A</td>
<td>31</td>
<td>$47.85</td>
</tr>
</tbody>
</table>
Figure 19 Our data collection and processing workflow for subsequent analysis in Section IV.

Steps and ⇤ are partially manual operations; the others are fully automated.

The rest of this section describes these steps in detail.

A. Collecting Spam-Advertised URLs

Our study is driven by a broad range of data sources of varying types, some of which are provided by third parties, while others we collect ourselves. Since the goal of this study is to decompose the spam ecosystem, it is natural that our seed data arises from spam email itself. More specifically, we focus on the URLs embedded within such email, since these are the vectors used to drive recipient traffic to particular Web sites. To support this goal, we collected feed data from August 1 through October 2012, which together comprised nearly 2 billion URLs.

Table I9 Feeds of spam-advertised URLs used in this study. We collected feed data from August 1 through October 2012.

<table>
<thead>
<tr>
<th>Feed Name</th>
<th>Description</th>
<th>URLs</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feed - MX honeypot</td>
<td>21437252</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td>Feed (Seeded honey accounts)</td>
<td>62513744</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Feed (Seeded honey accounts)</td>
<td>25137484</td>
<td>464</td>
<td></td>
</tr>
<tr>
<td>Feed X MX honeypot</td>
<td>88523618</td>
<td>181</td>
<td></td>
</tr>
<tr>
<td>Feed Y Human identified</td>
<td>62212424</td>
<td>224</td>
<td></td>
</tr>
<tr>
<td>Feed Z MX honeypot</td>
<td>46133656</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Grum</td>
<td>81338237</td>
<td>237</td>
<td></td>
</tr>
<tr>
<td>Mega</td>
<td>1114231</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Rustock</td>
<td>516251</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Other bots</td>
<td>657</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8578272</td>
<td>8641</td>
<td></td>
</tr>
</tbody>
</table>

Obtained seven distinct URL feeds from third-party partners, including multiple commercial antispam providers, and harvested URLs from our own botfarm environment. For this study, we used the data from these feeds from August 1 through October 2012, which together comprised nearly 2 billion URLs.

Note that the “bot” feeds tend to be focused spam sources, while the other feeds are spam sinks comprised of a blend of spam from a variety of sources. Further, individual feeds, particularly those gathered directly from botnets, can be heavily skewed in their makeup. For example, we received over 114 million URLs from the Grum bot, but these only contained 237 distinct registered domains. Conversely, the 224 million distinct domains produced by the Rustock bot are artifacts of a “blacklist poisoning” campaign undertaken by the bot operators that comprised millions of “garbage” domains.

Thus, one must be mindful of these issues when analyzing such feed data in aggregate.

From these feeds, we extract and normalize embedded URLs and insert them into a large multiterabyte PostgreSQL database. The resulting “feed tables” drive virtually all subsequent data gathering.

B. Crawler data

The URL feed data subsequently drives active crawling measurements that collect information about both the [NS infrastructure used to name the site being advertised and the Web hosting infrastructure that serves site content to visitors. We use distinct crawlers for each set of measurements.

DNS Crawler: We developed a [NS crawler to identify the name server infrastructure used to support spam-advertised domains and the address records they specify for hosting those names. Under normal use, this process would be straightforward, but in practice it is significantly more complex.

Levchenko et al., Click Trajectories: An End-to-End Analysis of the Spam Value Chain, 2011
- 120 items purchased
- 76 authorized
- 56 settled
- 49 products delivered

- 2 sent after mailbox lease ended
- 2 no follow-up email
- 2 resent after mailbox lease ended
- 1 promised refund (never obtained)

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Item</th>
<th>Origin</th>
<th>Affiliate Programs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aracoma Drug</td>
<td>Orange bottle of tablets (pharma)</td>
<td>WV, USA</td>
<td>ClFr</td>
</tr>
<tr>
<td>Combitic Global Caplet Pvt. Ltd.</td>
<td>Blister-packed tablets (pharma)</td>
<td>Delhi, India</td>
<td>GlvMd</td>
</tr>
<tr>
<td>M.K. Choudhary</td>
<td>Blister-packed tablets (pharma)</td>
<td>Thane, India</td>
<td>OLPPh</td>
</tr>
<tr>
<td>PPW</td>
<td>Blister-packed tablets (pharma)</td>
<td>Chennai, India</td>
<td>PhEx, Stmul, Trust, ClFr</td>
</tr>
<tr>
<td>K. Sekar</td>
<td>Blister-packed tablets (pharma)</td>
<td>Villupuram, India</td>
<td>WIdPh</td>
</tr>
<tr>
<td>Rhine Inc.</td>
<td>Blister-packed tablets (pharma)</td>
<td>Thane, India</td>
<td>RxPrm, DrgRev</td>
</tr>
<tr>
<td>Supreme Suppliers</td>
<td>Blister-packed tablets (pharma)</td>
<td>Mumbai, India</td>
<td>Eva</td>
</tr>
<tr>
<td>Chen Hua</td>
<td>Small white plastic bottles (herbal)</td>
<td>Jiangmen, China</td>
<td>Stud</td>
</tr>
<tr>
<td>Etech Media Ltd</td>
<td>Novelty-sized supplement (herbal)</td>
<td>Christchurch, NZ</td>
<td>Staln</td>
</tr>
<tr>
<td>Herbal Health Fulfillment Warehouse</td>
<td>White plastic bottle (herbal)</td>
<td>MA, USA</td>
<td>Eva</td>
</tr>
<tr>
<td>MK Sales</td>
<td>White plastic bottle (herbal)</td>
<td>WA, USA</td>
<td>GlvMd</td>
</tr>
<tr>
<td>Riverton, Utah shipper</td>
<td>White plastic bottle (herbal)</td>
<td>UT, USA</td>
<td>DrMax, Grow</td>
</tr>
<tr>
<td>Guo Zhonglei</td>
<td>Foam-wrapped replica watch</td>
<td>Baoding, China</td>
<td>Dstn, UltRp</td>
</tr>
</tbody>
</table>

Table VI: List of product suppliers and associated affiliate programs and/or store brands.
Companies already incorporated immediately available this advertisement of one typical provider. "We have ready-made shell route transactions through merchant accounts at difference cooperating third-party payment processors may be able to. But again, from the same set. This suggests that while funds and was promptly retried through two different banks. Order placed with rugRevenue failed due to insufficient zeroigazbankr on or around January 3rd. Finally, one xpressr all appear to have moved to this bank. A private commercial bank in Azerbaijan. &N 2zfi919r. appearing in our followvon purchases is bank Standard o S –ard Service of Germany. Indeed, the have rotated through two different Latvian banks and -Russiau while Royal Software (uroSoft and Soft Salesu Soft Store have started clearing through NOBank in pleu transactions with ZavPharmacyu [greenlineu and O(M. qtypically in January or February fiz3ru they still stayed Resellers. Moreover, while many programs did change from Online Pharmacy and all software from Softw herbal products sold through Zed–ashu all pharmaceuticals use the same banks four months later. ewgwu all replica and through the major affiliate programs. Many continued to four months after our study we continued to place orders alternatives and far higher switching cost. Value chain, we believe payment infrastructure has far fewer problems. Significant account "holdbacks" that they reclaim when there been unable to locate providers willing to provide operating merchant with both the bank and Visa Mastercard. We have a payment processor acts as middleman and "fronts" for the or weeks. Even for so-called third-party accounts whereby figure 30 Takedown effectiveness when considering domain registrars qleftru —NS and Web hosters qcenterr and acquiring banks qrightrw. Thus, unlike the other resources in the spam community, furthermore, for a subset of spam advertised banking resources—a rare asymmetry favoring the antivspam far more rapidly than the turnvaround time to acquire new by modest numbers of undercover buys, as in our study, and "financial blacklist" could be updated very quickly. Driven furthermore, it appears plausible that such a then the underlying enterprise would be dramatically dev the banks identified as supporting spam advertised goods with registrars [ ClickTrajectories: An End-to-End Analysis of the Spam Value Chain, 2011 ]
Can we throttle abuse by targeting merchant accounts at banks?

- McCoy et al., Priceless: The Role of Payments in Abuse-advertised Goods, 2012
- Made purchases to pharma and software OEM programs, while also working with brandholders to make complaints to Visa/MC
Figure 4: Example of a program receiving complaints to a card network. Rows denote distinct merchant descriptors; row “X” shows refused orders.

Wrote one eloquent affiliate in March of this year, “Right now most affiliate eprograms have a mass of declines, cancels and pendings, and it doesn’t depend much on the program IMHO, there is a general sad picture, f!#$@ing Visa is burning us with napalm.”

McCoy et al., Priceless: The Role of Payments in Abuse-advertised Goods, 2012
Ethics

- We have seen researchers:
  - measuring illicit activities of victims
  - participating in spam campaigns
  - taking ownership of bots / botnet C&C
  - purchasing goods from criminal organizations
  - port scanning victims

- Ethics discussion in papers:
  - short discussion justifying lack of harm
  - “beyond the scope of this work”

From paper on Torpig takeover (Stone-Gross et al.):

PRINCIPLE 1. The sinkholed botnet should be operated so that any harm and/or damage to victims and targets of attacks would be minimized.

PRINCIPLE 2. The sinkholed botnet should collect enough information to enable notification and remediation of affected parties.
Traffic sellers

• Click fraud
• Click traffic sellers
  – grey-market
  – Class project pilot study to see what these sellers are all about
    • Botnet traffic?
    • Legitimate project?
  – http://cseweb.ucsd.edu/~tristenp/buytraffic/
You can’t make sales if don’t have visitors

"30 days unlimited traffic"
Stop getting scammed from traffic sellers!
This is real quality traffic that
We use for own sites.

INCREASE WEB TRAFFIC GUARANTEED!
# Click traffic sellers

<table>
<thead>
<tr>
<th>Web site</th>
<th>CP10k</th>
<th>Claimed traffic source</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.trafficdeliver.com">www.trafficdeliver.com</a></td>
<td>~$34.69</td>
<td>“Advertiser exchange”</td>
</tr>
<tr>
<td>revisitors.com</td>
<td>~$48.95</td>
<td>Recently expired domain redirection?</td>
</tr>
<tr>
<td>qualitytrafficsupply.com</td>
<td>~$55.00</td>
<td>Contextual advertisements</td>
</tr>
<tr>
<td>mediatraffic.com</td>
<td>~$70</td>
<td>AdWare (Voomba) pop-ups</td>
</tr>
</tbody>
</table>

**Targeted vs. untargeted:** specify geographic preferences

**Affiliate networks:** paid to send traffic

**Traffic resellers:** resell purchased traffic
Experimental methodology

(1) Setup several web sites (xxx.sysnet.ucsd.edu)
   2 pages: index.html is landing site
   lucky.html linked to by index.html

   Example site linked from webpage

(2) Attempt to purchase web traffic

   Used temporary VISA number, but real name, etc.

(3) Sit back and let the research data come to us ...
Adventures in purchasing web traffic...

Giving people money not as easy as I expected:

- revisitors.com
- qualitytrafficsupply.com
- mediatraffic.com
- www.trafficdeliver.com

Took my money
Sent "targeted" US traffic
Took my money
No response...
Wanted $200 deposit
Took my money ...
... but gave it back!

---

RE: Refund - [2423-DLXC-4301] [82a2e44b]

2Checkout Help Desk  ========= Please enter your reply ABOVE above this line ========= Hello Tom, ...

2Checkout Help Desk  A staff member has replied to your question: Seasons Greetings Tom, Thank you...

2Checkout Help Desk  Thank you for adding a message to your question. We will respond to your mess...

2Checkout Help Desk  to me

======== Please enter your reply ABOVE above this line ========

Hello Tom,

A staff member has replied to your question:

Dear Tom,

Thank you for contacting 2Checkout.com. I apologize for the delay in responding to your inquiry. The order was actually canceled trafficdeliver.com. They believe the order to be fraudulent. I have forwarded your inquiry to trafficdeliver.com. They will be contacting you via e-mail shortly. If you do not receive a response in a timely manner, please feel free to reopen this ticket for additional assistance.

Looking to make your holidays happier? 2Checkout makes it easy! Simply visit your favorite search engine and type in 2Checkout + and the type of merchandise you are looking for. It’s the easy way to enjoy a fast, safe shopping experience online.

Thank You,
Josh Karamian
Customer Care
2Checkout.com
http://www.2Checkout.com

Dec 6 (5 days ago)
Dec 6 (5 days ago)
Dec 6 (4 days ago)
When did traffic arrive?
When did traffic arrive?

- Not a typical pattern for traffic
When did traffic arrive?

- Traffic has really high-degree of temporal proximity
- Anecdote: many IPs visit times clustered within seconds
Is the traffic from bots or other malware?

<table>
<thead>
<tr>
<th>Source</th>
<th>Num IPs</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBL</td>
<td>21</td>
<td>1.7%</td>
</tr>
<tr>
<td>Current Storm</td>
<td>0</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Other interesting anecdotal evidence

4 HEAD requests from distinct IPs with referrer

http://www.routetraffic.net/delivery/statistics/8x0ada67md29fk799sa4.html
Next we examine the distributions of User-Agent strings of the vendor employs dubious means for delivering traffic to their clients. With one (now-defunct) bulk traffic vendor QualityTrafficSupply, such behavior serves as a heavy-handed signature that the vendor is delivering traffic to their clients. Although annoying to deal with, such behavior serves as a heavy-handed signature that the vendor is delivering traffic to their clients.

Using our snapshotting tool to visit the referrers induced a HTTP denial-of-service attack on our server. Although annoying to deal with, such behavior serves as a heavy-handed signature that the vendor is delivering traffic to their clients.

To avoid this and conserve space, we enabled our snapshotting tool to automatically snapshot the referring page. To avoid this and conserve space, we enabled our snapshotting tool to automatically snapshot the referring page.

We use the User-Agent field in HTTP requests, when present, to identify the browser and operating system on which the browser is run. In this case, we use the User-Agent field in HTTP requests, when present, to identify the browser and operating system on which the browser is run.

Visitors to our sites as another possible signature. The User-Agent field in HTTP requests identifies the client software used to make the request. Web browsers set the field to identify the browser software and the operating system on which the browser is run.

Visitors to our sites as another possible signature. The User-Agent field in HTTP requests identifies the client software used to make the request. Web browsers set the field to identify the browser software and the operating system on which the browser is run.

Visitors to our sites as another possible signature. The User-Agent field in HTTP requests identifies the client software used to make the request. Web browsers set the field to identify the browser software and the operating system on which the browser is run.

Visitors to our sites as another possible signature. The User-Agent field in HTTP requests identifies the client software used to make the request. Web browsers set the field to identify the browser software and the operating system on which the browser is run.

Visitors to our sites as another possible signature. The User-Agent field in HTTP requests identifies the client software used to make the request. Web browsers set the field to identify the browser software and the operating system on which the browser is run.
Attention: Cs.wisc.edu Web User,
You have exceeded your e-mail account limit quota of 250MB and you are requested to expand it within 48 hours or else your e-mail account will be disable from our database. Simply CLICK HERE <https://docs.google.com/spreadsheet/viewform?formkey=dERrcTlFQ2tFZ3hETkkzcVc1UjMxWmc6MQ> with the complete information requested to expand your e-mail account quota to 450MB.
Thank you for using indonet e-mail services.
Copyright ©2012  cs.wisc.edu Information Center.
E-crime is a complex ecosystem

- Lots of moving parts
- Economics important
  - Fascinating measurement studies
- Technical mechanisms often don’t measure up
- “In Planning Digital Defenses, the Biggest Obstacle Is Human Ingenuity” - Stefan Savage